

## 論文の内容の要旨

論文題目           Robust Speech Recognition for Handheld Devices

(携帯機器向け耐雑音音声認識の研究)

氏   名           大 淵           康 成

音声認識は、いまやPDAや携帯電話などの携帯機器においても使用可能になりつつある。以前はハードウェアの性能の制約により不可能だったが、マイクロプロセッサ技術の革新により、これらのような小さな機器に音声認識機能を持たせることが現実味を帯びつつある。そして言うまでもなく、キーボードを持たないこれらの機器において、音声認識機能の重要性は極めて高い。

携帯機器における音声認識機能が現実的になってきたとはいえ、いくつかの課題が残されていることは間違いない。どのような場面で、どのような使い方で携帯機器が用いられるかに応じて、様々な困難が生じることは想像に難くない。そしてそれらの問題は、単一の解ですべて解決されるものでもない。

本論文の目的は、携帯機器に音声認識を実装する際の諸問題を解析し、その実現のための明確なガイドラインを提供することにある。加えて、その実現のために必要となる新しいアルゴリズムを示すことも重要な目的である。

音声認識システムは、複数のサブシステムの組合せとして構成される。音響モデルや言語モデルの学習など、いくつかのサブシステムに関しては、一般的なコンピュータ上で用いられている方式をそのまま使用することが可能である。そのように課題を分割して考えると、コンパクト性とロバスト性という、携帯機器においてとりわけ重要な二つの課題が最終的に明らかになる。

携帯機器においては、メモリやCPUなどの計算資源が限られている。速やかな反応を得るためには、アルゴリズムは十分に高速でなければならない。実際、これまでも音声認識アルゴリズム高速化の様々な方式が提案されているが、ハードウェアが未成熟だったこともあり、実際の製品として使用可能な形でこれら諸々のアルゴリズムが統合された例は無かった。しかし、そのような時を経て、ようやく必要な技術が得られるようになった。我々の研究はそこから始まり、携帯機器向けのマイクロプロセッサに焦点を絞ったうえで、様々な技術を統合・最適化し、音声認識ミドルウェアという形の成果を得た。このミドルウェアは、多くの製品において実際に使用されている。

次の段階として、我々はロバスト性の問題に着目した。携帯機器は屋外を含めた様々な環境で用いられることが多く、音声入力が必要な環境雑音により妨害されるという前提のもとで、頑強なアルゴリズムを開発しなければならない。そこで我々はデルタケプストラム正規化(DCN)という新アルゴリズムを提案する。DCNは、わずかな計算量の増加により、雑音環境下での頑強性を実現するアルゴリズムである。DCNの提案は本論文の主要な成果のひとつであり、他の部分においても多く言及されることから、その数理的定式化、可変パラメータの最適化、および様々な環境における実験を通じての評価を行ない、DCNの詳細を明らかにする。

第三段階では、音声認識の更なる性能向上をターゲットとする。これまでの成果によって、携帯機器における音声認識を実現することはできたが、更に劣悪な環境を考えた場合、アルゴリズムの頑強性はいまだ十分とは言えない。このような劣悪な環境に対応することは困難な課題であるが、いくつかのケースにおいては、速度の制約を緩めて考えることが可能であり、そのような場合には、いわゆるマイクロフォンアレイの枠組みが有効である。そこで、携帯機器に取り付けた複数のマイクロフォンの特性を調べる。その結果として得られる携帯機器に特有の性質を解析した結果、チャンネル選択と特徴量合成という二つのアプローチが有効であることが明らかとなった。本論文では、それぞれのアプローチに対して新しいアルゴリズムを提案し、各々が従来のマイクロフォンアレイ手法よりも高い性能を持つことを示す。

これらの成果により、携帯機器における音声認識の主要な問題点が明らかとなるとともに、新規に提案したアルゴリズムの活用によって音声認識の性能が向上することが示された。さらにこれらの結果により、携帯機器における音声認識の更なる性能向上へ向けて、明確な方向付けを与えることができた。